DTIC FILE COPY (2)

MUMPS BASED INTEGRATION OF DISPARATE COMPUTER-ASSISTED MEDICAL DIAGNOSIS MODULES

A. GINO
W. M. PUGH
D. H. RYMAN

REPORT NO. 89-43

NAVAL HEALTH RESEARCH CENTER
P.O. BOX 85122
SAN DIEGO, CALIFORNIA 92186-5122

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND
BETHESDA, MARYLAND

MUMPS BASED INTEGRATION OF DISPARATE COMPUTER ASSISTED MEDICAL DIAGNOSIS MODULES

Antonio Gino, Ph.D.
William M. Pugh
Dave H. Ryman

Medical Decision Support Department
Naval Health Research Center
P.O. Box 85122
San Diego, CA 92138-9174

*Report 89-43, supported by the Naval Research and Development Command, Department of the Navy, under Work Unit No. M0095.005-6051. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense, nor the U.S. government. Approved for public release, distribution unlimited.

SUMMARY

Two problems are frequently faced in the application of Artificial Intelligence ~~(AI)~~ techniques to Computer-Assisted Medical Diagnosis (CAMD): (a) inability to use different techniques within the same CAMD system, and (b) difficulty of sharing data between the CAMD and the patient data base. Modules that use Bayesian computations to estimate the likelihood of certain diagnoses, for example, are kept separate from modules that use a Rule or a Neural Network Model based approach. Over the past several years, the U.S. Navy has funded development of a number of CAMD modules, however, the modules are not uniform with respect to the process used to arrive at the diagnosis. For example, the Abdominal and Chest Pain modules use a Bayesian approach, while the Opthalmology module uses a Rule Based approach. In the current effort, MUMPS is used to develop an integrated, easily expandable CAMD system capable of executing virtually any AI technique (Bayesian, Rule Based, Neural Network Model based, etc.). At the same time, MUMPS is used to create an integrated medical/diagnostic encounter data base. (KR)

MUMPS-BASED INTEGRATION OF DISARATE COMPUTER-ASSISTED MEDICAL DIAGNOSIS MODULES

Introduction

The absence of medical doctors in the medical departments of many U.S. Navy ships requires the medical services to be rendered by hospital corpsmen. To assist the corpsmen, the U.S. Navy has initiated a program to provide them Computer Assisted Medical Diagnosis (CAMD). This effort has resulted in the development of a number of modules: Abdominal Pain, Chest Pain, Dental Pain, Psychiatric, and Ophthalmology. The final system will allow corpsmen in isolated environments (submarines, surface ships, or isolated bases) to arrive at a differential diagnosis or manage an injury, i.e., decide whether a disease or injury warrants evacuation and determine what preliminary care can be given to the patient during the preparation for evacuation.

The development of CAMD modules has traditionally followed one of two paths: Bayesian or Expert System. In a Bayesian CAMD module, the prior probabilities of the diseases to be diagnosed and the conditional probabilities of symptoms are used to arrive at the diagnoses, given the signs and symptoms reported by the patient [1, 6, 8, 9, 10]. In an Expert (or Rule Based) System, logical steps are used to identify the most probable disease(s) for the patient, i.e., If-Then rules are combined to arrive at a logical ordering of the possible diagnoses given the signs and symptoms [3, 4, 5, 7].

More recently, a third approach, the Neural Network Model, has been used [12, 13, 14]. This approach involves the emulation of the information processing by neurons. Undoubtedly, the number of techniques used for CAMD development will continue to grow and this raises the issue of integrating all the modules developed. For example, the modules developed by the Navy use Baysian (Abdominal and Chest Pain modules) and Expert System (Psychiatric and Ophthalmology modules) approaches to estimate the likelihood of disease. The modules are also "stand alone". That is, they lack a common data dictionary and have different user interfaces. The result is a fragmented data base and the need for different training for the various modules since the user interface is not the same across modules. This obviously creates a host of problems. From the perspective of the user, any information about the patient entered for, say, the Abdominal module, will not be available if the user accesses the other modules. Therefore, overlapping areas (e.g., Abdominal and Chest Pain) cannot be simultaneously compared. From the perspective of system management, since there is no unified approach to development, new modules cannot be easily integrated into an existing structure. Finally, the modules do not use the same programming language, which requires expertise in a number of areas for maintenance. This paper describes our approach to the creation of an integrated system from the various "stand alone" modules already developed. Because a multi-purpose "Shell" is at the heart of the approach taken, the shell is described in detail and many potential uses are highlighted.

A General "Shell" Solution

Over the past year a "shell" has been developed that is expected to solve the problem of executing Bayesian, Neural Network or Rule based CAMDs from one system. Figure 1 shows a schematic of the system we have developed. As one can see from the figure, we have one central driver for all systems (CAMD Main Driver), an integrated patient data base (Patient File), and Baysian, Rule Based, and Neural Network model drivers.

A Data Base Management Approach (DBMS) was taken to resolve two issues raised above (integration among systems and integration between systems and patient data). Using the VA FileManager, a CAMD System File, shown in Table 1, has been defined that contains the Data Dictionary for the modules (Abdominal Pain, Chest Pain, etc.). A set of programs that implement the Bayesian and Neural Net computations used to determine a diagnosis has also been developed. The programs obtain instructions from the CAMD System file and carry out the computations.

Fig.1. CAMD FLOW DIAGRAM

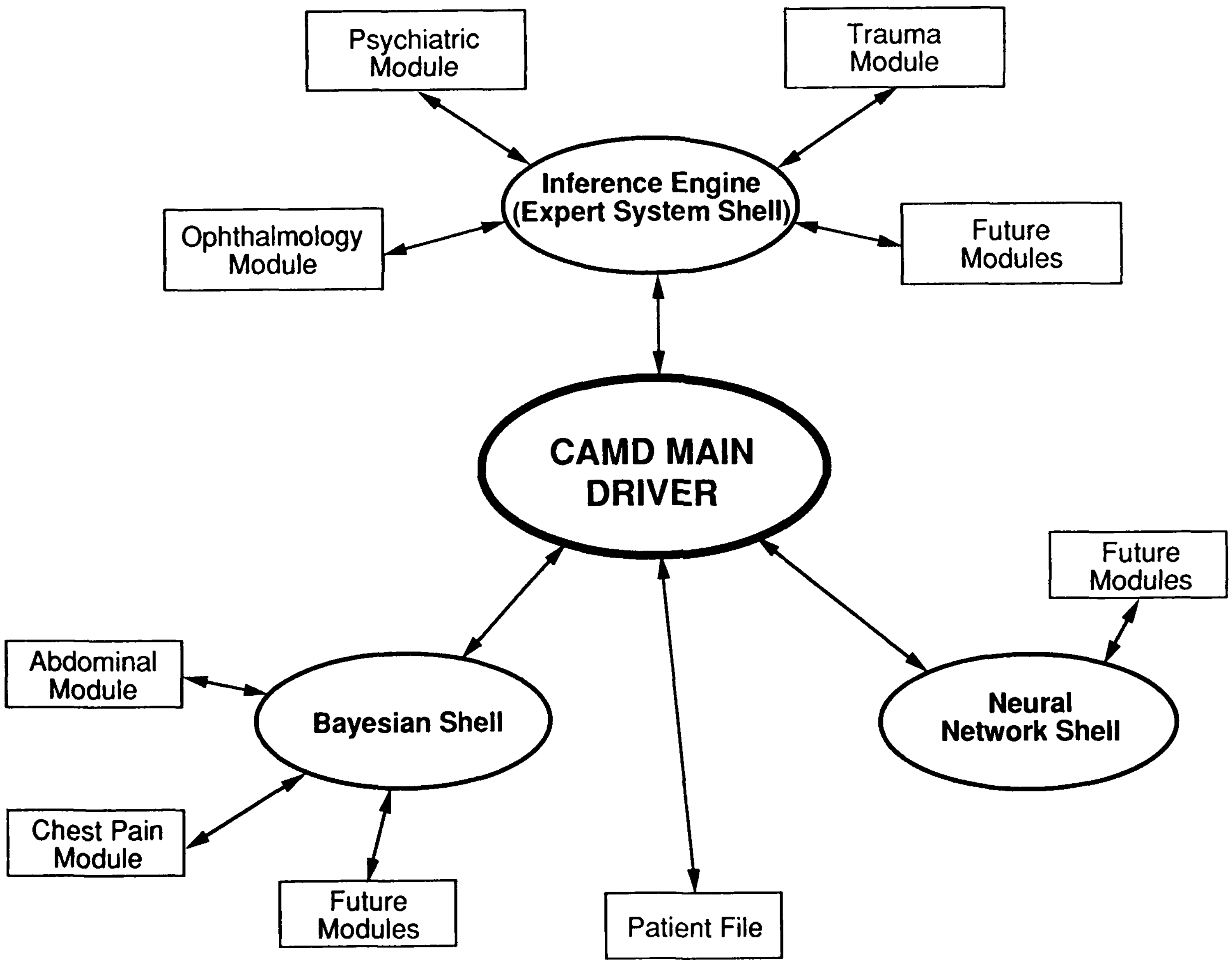

Table 1. CAMD SYSTEM FILE

Files pointed to (arrows) from inside the CAMD System File are also shown to give a more complete picture of the structure of the system.

.01	SYSTEM NAME	(e.g., ABDOMINAL)
.02	SYSTEM TYPE	(e.g., Bayesian)
1	BIAS	(Neural Network)
1.1	INTERMEDIATE LAYERS	(Neural Network)
10	INPUTS	(possible symptoms, multiple)
.01	Symptom	
.02	Weight	(Neural Network)
.03	A Priori Probability	(Bayesian)
20	OUTPUTS	(possible diseases, multiple)
.01	Disease	
.02	A priori Probability	(Bayesian)
30	LAYER	(Neural Net, multiple)
.01	Layer Number	(1, 2, etc.)
10	NODE	(multiple)
.01	Node Number	(1, 2, etc.)
.50	Weight Assigned	(computed during training)

SYMPTOM FILE

.01	SYMPTOM NAME
.02	Subsystem (e.g., Abdominal)

DISEASE FILE

.01	DISEASE NAME
10	TREATMENT (Word Processing)
20	SIMILAR DISEASES
.01	Similar Disease
30	DISCRIMINATING SYMPTOMS
.01	Symptom

Assume we wanted to incorporate a Bayesian CAMD module into our "shell". The (simplified) procedural steps are as follows:

1. Obtain cases needed to compute the *a priori* probabilities of diseases and symptoms.
2. Compute the weights associated with each symptom.
3. Extract the weights from the Bayesian system.
4. Enter the Bayesian weights into the "shell" (i.e., add the symptoms and diseases, with associated *a priori* probabilities into the System File).
5. Add the Bayesian module to the overall menu.

A Neural Network Model based CAMD can be incorporated via an entirely analogous procedure (the system is designed to handle only back propagation type networks with a sigmoidal activation function, [15, 16]):

1. Obtain cases needed to "train" the neural network.
2. Train the neural net until the weights stabilize.
3. Extract the weights from the neural net.
4. Enter the weights obtained from the neural net into the "shell".
5. Add the Neural Network module to the overall menu.

The example in Table 2 clarifies the process. Assume that we have a CAMD developed with a Neural Network Model with three inputs (symptoms), one intermediate layer (with two nodes), and two possible outputs (diseases). We would need the following file layout to simulate the (predictive) behavior of the network within our "shell" (fields that are not shown are assumed to contain nothing).

Table 2. CAMD EXAMPLE

Field Name	Contents
SYSTEM NAME	EXAMPLE
SYSTEM TYPE	Neural Net
BIAS	.5
INTERMEDIATE LAYERS (NUMBER)	1
INPUT	
Symptom	1
Weight Assigned	-0.8
Symptom	2
Weight Assigned	+0.3
Symptom	3
Weight Assigned	+0.9
OUTPUT	
DISEASE	1
DISEASE	2
LAYER	
Layer Number	1
NODE	
Node Number	1
Weight Assigned	+0.5
Node Number	2
Weight Assigned	-0.3

A Rule Based system requires additional considerations. The objective is to extract the rules from a Rule Based system, transfer them to the CAMD shell, and combine them with the patient's symptoms to obtain the likelihood of diseases. Therefore, a general procedure is being developed to allow communication between the CAMD shell and any external Expert System Shell (ESS) loaded with rules. The assumption made is that the ESS can be called from the shell and can return answers (possible diagnoses) in a standard format. That is, one should be able to pass data collected within our environment (shell) to the ESS (either as an ASCII file or as a MUMPS global), and obtain the diagnoses from the ESS (again, either as an ASCII file or as a MUMPS global). Currently, the ESS used is RuleMan, a MUMPS based application developed by the Veterans Administration [2].

Incorporation of a Rule Based CAMD requires the following specifications:

1. ASCII file or MUMPS global needed for input (to the ESS) and output (from the ESS).
2. Format for each input (symptom) and output (diagnosis/disease).

For example, the ESS output may consist of an ASCII file where the most likely diagnosis is stored in the first line of the file, the second most likely in the second line, etc. Furthermore, each diagnosis may be identified by a number, and the number may be stored in specific columns, etc. Analogously, the ESS may require that the first symptom reported be stored in the first line of an ASCII file.

Summary of Steps Executed by the "Shell"

The "shell" described is currently under development (although portions have been completed). The final product is expected to execute the following steps:

1. Obtain information from available data bases [e.g., demographics from the Shipboard Non-tactical ADP Program (SNAP) Automated Medical System (SAMS), Composite Health Care System (CHCS), etc.] and from the user (signs and symptoms).

2. Pass the information obtained to either a program that computes the "likelihood" of a number of diagnoses given the symptoms (e.g., the program applies Bayesian, Neural Net, regression, etc. weights); or the Inference Engine of an Expert System (where the rules for the various diagnoses have been stored).

3. Obtain the "answers" from #2 above and display the possible diagnoses to the user.

4. Loop back to #1 until the user is satisfied. That is, obtain additional information from the user and re-compute the possible diagnoses (re-run the programs) until the user indicates he/she does not wish to continue the process (or selects a diagnosis).

5. Give, at the user's discretion, treatment recommendations for any diagnoses that have been found to be a possibility, and the methods by which the diagnosis was reached.

6. Store any relevant information into the patient file (e.g., the date/time of the encounter between the user and the CAMD system, all data/symptoms provided by the user or obtained from data bases, all diagnoses inferred, treatment recommendations given, etc.).

Discussion

The integrated solution being pursued has many advantages. Since it is a "shell" approach, it will greatly facilitate the incorporation of new CAMDs into an overall CAMD system. Since it is a MUMPS solution, interfaces with the many medical data bases within the Department of Defense (DOD) and the Veterans Administration (VA) can be easily developed. Integration between our CAMD "shell" and the VA or DoD medical software is rather straightforward, at least as far as the strictly MUMPS portions of the software are concerned (see below). In addition, since we are creating a patient encounter data base, integration of our patient data with other MUMPS medical data bases should be facilitated.

The major shortcoming of the MUMPS solution is its possible lack of transportability, which is counter to what one would expect, given that transportability of software is an often touted feature of the MUMPS language. The shortcoming is due to our need to use the ZCALL feature, the only way to execute code written in other computer languages (expert systems). Since the ZCALL is, unfortunately, not yet one of the standard features of MUMPS, we may have to develop separate ZCALL utilities for each separate MUMPS implementation.

References

1. DeDombal, F.T., Leaper, D.J., Staniland, J.R., McCann, A.P., & Horrocks, J.C., Computer-Aided Diagnosis of Acute Abdominal Pain. British Medical Journal, Vol. II, 9-13, 1972.

2. Dymond, A.M., An Expert System for the DHCP Environment. MUG Quarterly, Vol. XVIII, No. 1, 6-9.

3. Elion, J.L., Expert Advisor System for the Management of Interventions in Acute Myocardial Infraction. IEEE Annual Meeting of Computers in Cardiology, Sept 25-28 1988, Washington D.C.

4. Hickman, D.H., Shortliffe E.H., Bischoff, M.S., Scott, A.C., Jacobs, C.D., The Treatment Advice of a Cancer Chemotherapy Protocol Advisor. Ann Intern Medicine, 103, 928-936, 1985.

5. Kingsland, L.C. III, Lindberg, D.A.B. and Sharp, G.C., Anatomy of Knowledge-based Consultant System: AI/Rheum. MD Computing, 3, 18-26, 1986.

6. Ledley, R.S. and Lusted, L.B., Reasoning Foundations of Medical Diagnosis. Science , 130, 9-12, 1959.

7. Miller R.A., Pople H.E., Meyers J.D., INTERNIST-I, An Experimental Computer-based Diagnostic Consultant for General Internal Medicine. New England Journal of Medicine, 307, 468-476, 1982.

8. Reggia, J.A., Tabb, D.R., Price, T.R., Banko, M. and Hebel, R.: Computer-aided Assessment of Transient Ischemic Attacks. Arch Neurol (1984), 41, 1248-1255.

9. Sherman, H., A Pocket Diagnostic Calculator Program for Computing Bayesian Probabilities for Nine Diseases with Sixteen Symptoms. Comp Biomed Res (1978) 11, 177.

10. Warner, H.R., Toronto, A.F. and Veasy, L.G., Experience with Bayes Theorem or Computer Diagnosis of Congenital Heart Disease. Ann N Y Acad Sci (1964), 115, 2-26.

11. Slagle, J.R., Applications of Generalized Network-based Expert System Shell. Twelfth Annual Symposium on Computer Applications in Medical Care, Washington, D.C., pp. 33-42, 1988.

12. Mulsant, B.H. & Servan-Schreiber, E., A Connectionist Approach to the Diagnosis of Dementia. Twelfth Annual Symposium on Computer Applications in Medical Care, Washington, D.C., pp. 245-250, 1988.

13. Hudson, D.L., Cohen, M.E., & Anderson, M.F., Determination of Testing Efficacy in Carcinoma of the Lung using a Neural Network Model. Twelfth Annual Symposium on Computer Applications in Medical Care, Washington, D.C., pp. 251-255, 1988.

14. Hripsak, G., A Neural Network for Medical Diagnosis that Accomodates Unreliable Data. Proceedings of the AAMSI Congress 89, San Francisco, California, pp. 159-163, 1989

15. Rumelhart, D.E. & McClelland, J.L. (Eds.): Parallel Distributed Processing: Vol. 1, Foundations. Cambridge, MA., 1986.

16. Lippman, R.P. An Introduction to Computing with Neural Networks. IEEE ASSP Magazine, pp. 4-22, April 1987.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS NONE
2a SECURITY CLASSIFICATION AUTHORITY N/A 2b DECLASSIFICATION / DOWNGRADING SCHEDULE N/A	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NHRC Report No. 89- 43	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Health Research Center 6b OFFICE SYMBOL (If applicable) Code 20	7a NAME OF MONITORING ORGANIZATION Commander Bureau of Medicine and Surgery
6c ADDRESS (City, State, and ZIP Code) P.O. Box 85122 San Diego, CA 92138-9174	7b ADDRESS (City, State, and ZIP Code) Department of the Navy Washington, DC 20372
8a NAME OF FUNDING SPONSORING ORGANIZATION Navy Medical Research and Development Command 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code) NNMC Bethesda, MD 20814-5044	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
63706N	M0095	005	6051

11 TITLE (Include Security Classification)

MUMPS BASED INTEGRATION OF DISPARATE COMPUTER ASSISTED MEDICAL DIAGNOSIS MODULES

12 PERSONAL AUTHOR(S)
GINO, A., PUGH, W.M., RYMAN, D.H.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM ______ TO ______	1989 DECEMBER 12	

16 SUPPLEMENTARY NOTATION

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Computer-Assisted Medical Diagnosis, Bayesian, Rule-Based, Neural Network Integration

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Two problems are frequently faced in the application of Artificial Intelligence (AI) techniques to Computer-Assisted Medical Diagnosis (CAMD): (a) Inability to use different techniques within the same CAMD system and (b) Difficulty of sharing data between the CAMD and the medical, patient data base. Modules that use Bayesian computations to estimate the likelihood of certain diagnoses, for example, are kept separate from modules that use a Rule or Neural Network Model based approach. Over the past several years, the U.S. Navy has funded development of a number of CAMD modules (Abdominal, Chest Pain, Dental, etc.,). The modules are not uniform with respect to the process used to arrive at the diagnosis, e.g., the Abdominal and Chest Pain modules use a Bayesian approach, while the Opthalmology module uses a Rule Based approach. In the current effort, MUMPS is used to develop an integrated, easily expandable CAMD system capable of executing virtually any AI technique (Bayesian, Rule Based, Neural Network Model based, etc.,). At the same time, MUMPS is used to create an integrated medical/diagnostic encounter data base.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Antonio Gino, Ph.D.	22b TELEPHONE (Include Area Code) (619) 553-8393 22c OFFICE SYMBOL Code 20

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

*U.S. Government Printing Office: 1985—507-047